

RECEIVED

MAR 04 2002

TECH CENTER 1600/2900



OIIPE

## RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/884,566A

TIME: 09:27:02

Input Set : A:\37497.txt

Output Set: N:\CRF3\02142002\I884566A.raw

ENTERED

## SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Keegan, Kathleen S.

8 (ii) TITLE OF INVENTION: Novel CREBa Isoform

10 (iii) NUMBER OF SEQUENCES: 10

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Marshall, Gerstein & Borun

14 (B) STREET: 233 South Wacker Drvie, 6300 Sears Tower

15 (C) CITY: Chicago

16 (D) STATE: Illinois

17 (E) COUNTRY: USA

18 (F) ZIP: 60606

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/884,566A

C--> 28 (B) FILING DATE: 19-Jun-2001

29 (C) CLASSIFICATION:

31 (viii) ATTORNEY/AGENT INFORMATION:

32 (A) NAME: Williams Jr., Joseph A.

33 (B) REGISTRATION NUMBER: 38,659

34 (C) REFERENCE/DOCKET NUMBER: 27866/37497

36 (ix) TELECOMMUNICATION INFORMATION:

37 (A) TELEPHONE: 312-474-6300

38 (B) TELEFAX: 312-474-0448

41 (2) INFORMATION FOR SEQ ID NO: 1:

43 (i) SEQUENCE CHARACTERISTICS:

44 (A) LENGTH: 3190 base pairs

45 (B) TYPE: nucleic acid

46 (C) STRANDEDNESS: single

47 (D) TOPOLOGY: linear

49 (ii) MOLECULE TYPE: cDNA

52 (ix) FEATURE:

53 (A) NAME/KEY: CDS

54 (B) LOCATION: 304..1866

57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

59 GGCACGAGGG ACTTCTTGG GATGAGCGCT GCCTTTTGG CTTCTTTTG GATGCACAGC	60
61 CCGATTTAAC CCCTGCACCT TCCGCCGAT CCCAGCAGGC TTGTCCTCCC CGGGGAGTCA	120
63 CAGATTTCCG AGGACAAGGG TCGCGTAGCC TTCGGCAGGG CTCTCCCGAG TTCCTGCTCC	180
65 AGTGCATAAG TTCCACGCGC GCACACGCCA AGTACACGGG GAGAAGCGTC TCACCGGCC	240

## RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/884,566A

TIME: 09:27:02

Input Set : A:\37497.txt

Output Set: N:\CRF3\02142002\I884566A.raw

67	GCGGCGGCTC	TGCGCGGTCC	CCTCCTGCCT	CAGCATCCTC	GGGCCTGCGC	GGCGCCACCC	300
69	GCC	ATG	GAG	GTG	CTG	GAG	348
70	Met	Glu	Val	Leu	Glu	Ser	
71	1		5		10	15	
73	CGC	AAG	CTG	AGC	GAG	CTG	396
74	Arg	Lys	Leu	Ser	Glu	Leu	
75		20		25		30	
77	TAC	CAC	ACG	CAC	TTC	TCG	444
78	Tyr	His	Thr	His	Phe	Ser	
79		35		40		45	
81	CTG	GGT	CAG	CTC	CTG	AGT	492
82	Leu	Gly	Gln	Leu	Leu	Ser	
83		50		55		60	
85	ATG	GAG	GTG	GAG	CCA	TCT	540
86	Met	Glu	Val	Glu	Pro	Ser	
87		65		70		75	
89	GAA	CAC	AGC	TAC	TCT	CTG	588
90	Glu	His	Ser	Tyr	Ser	Leu	
91	80		85		90	95	
93	ACC	CAT	GCG	GCT	ACC	AGC	636
94	Thr	His	Ala	Ala	Thr	Ser	
95		100		105		110	
97	GAA	AAA	TGG	TAC	CTG	TCT	684
98	Glu	Lys	Trp	Tyr	Leu	Ser	
99		115		120		125	
101	GAG	CCA	ATC	ACA	GAG	GAG	732
102	Glu	Pro	Ile	Thr	Glu	Glu	
103		130		135		140	
105	CTG	ACC	ATC	ACA	GCC	ATT	780
106	Leu	Thr	Ile	Thr	Ala	Ile	
107		145		150		155	
109	CTG	GAT	ATG	AAT	GCT	GGG	828
110	Leu	Asp	Met	Asn	Ala	Gly	
111	160		165		170	175	
113	AAG	ATT	AAG	CTG	GAG	CCC	876
114	Lys	Ile	Lys	Leu	Glu	Pro	
115		180		185		190	
117	CCG	AAA	GAA	GCC	TCC	GTG	924
118	Pro	Lys	Glu	Ala	Ser	Val	
119		195		200		205	
121	AGT	AGT	CAC	AGC	AGT	GAC	972
122	Ser	Ser	His	Ser	Ser	Asp	
123		210		215		220	
125	CTG	CAT	CCC	TTC	AGC	CTG	1020
126	Leu	His	Pro	Phe	Ser	Leu	
127		225		230		235	
129	CCC	CGG	GGC	CCC	TCT	GCC	1068
130	Pro	Arg	Gly	Pro	Ser	Ala	
131	240		245		250	255	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,566A

DATE: 02/14/2002

TIME: 09:27:02

Input Set : A:\37497.txt

Output Set: N:\CRF3\02142002\I884566A.raw

133	CAT	AAG	CTG	CAG	GGA	TCG	GGC	CCC	CTG	GTC	CTG	ACA	GAA	GAG	GAG	AAG	1116
134	His	Lys	Leu	Gln	Gly	Ser	Gly	Pro	Leu	Val	Leu	Thr	Glu	Glu	Glu	Lys	
135				260						265					270		
137	AGG	ACC	CTG	GTT	GCC	GAG	GGC	TAT	CCC	ATT	CCC	ACC	AAG	CTG	CCT	CTG	1164
138	Arg	Thr	Leu	Val	Ala	Glu	Gly	Tyr	Pro	Ile	Pro	Thr	Lys	Leu	Pro	Leu	
139				275					280					285			
141	ACA	AAA	TCT	GAG	GAG	AAG	GCC	CTG	AAG	AAA	ATC	CGG	AGA	AAG	ATC	AAG	1212
142	Thr	Lys	Ser	Glu	Glu	Lys	Ala	Leu	Lys	Lys	Ile	Arg	Arg	Lys	Ile	Lys	
143				290				295						300			
145	AAT	AAG	ATT	TCT	GCC	CAA	GAA	AGC	AGA	AAG	AAG	AAA	GAA	TAC	ATG		1260
146	Asn	Lys	Ile	Ser	Ala	Gln	Glu	Ser	Arg	Arg	Lys	Lys	Lys	Glu	Tyr	Met	
147		305					310					315					
149	GAC	AGC	CTG	GAG	AAA	AAA	GTG	GAG	TCT	TGT	TCA	ACT	GAG	AAC	TTG	GAG	1308
150	Asp	Ser	Leu	Glu	Lys	Lys	Val	Glu	Ser	Cys	Ser	Thr	Glu	Asn	Leu	Glu	
151	320				325					330					335		
153	CTT	CGG	AAG	AAG	GTG	GAG	GTG	CTG	GAG	AAC	ACC	AAT	AGG	ACT	CTC	CTT	1356
154	Leu	Arg	Lys	Lys	Val	Glu	Val	Leu	Glu	Asn	Thr	Asn	Arg	Thr	Leu	Leu	
155				340					345					350			
157	CAG	CAA	CTT	CAG	AAG	CTT	CAG	ACT	TTG	GTG	ATG	GGG	AAG	GTC	TCT	CGA	1404
158	Gln	Gln	Leu	Gln	Lys	Leu	Gln	Thr	Leu	Val	Met	Gly	Lys	Val	Ser	Arg	
159				355				360						365			
161	ACC	TGC	AAG	TTA	GCT	GGC	ACA	CAG	ACT	GGC	ACC	TGC	CTC	ATG	GTC	GTT	1452
162	Thr	Cys	Lys	Leu	Ala	Gly	Thr	Gln	Thr	Gly	Thr	Cys	Leu	Met	Val	Val	
163			370				375					380					
165	GTG	CTT	TGC	TTT	GCT	GTT	GCA	TTT	GGA	AGC	TTC	TTT	CAA	GGC	TAT	GGG	1500
166	Val	Leu	Cys	Phe	Ala	Val	Ala	Phe	Gly	Ser	Phe	Phe	Gln	Gly	Tyr	Gly	
167		385					390					395					
169	CCT	TAT	CCT	TCT	GCC	ACC	AAG	ATG	GCT	CTG	CCC	AGC	CAG	CAT	CCT	CTG	1548
170	Pro	Tyr	Pro	Ser	Ala	Thr	Lys	Met	Ala	Leu	Pro	Ser	Gln	His	Pro	Leu	
171	400				405					410					415		
173	TCA	GAG	CCA	TAC	ACA	GCC	TCC	GTG	GTG	AGA	TCC	AGG	AAC	CTG	CTA	ATC	1596
174	Ser	Glu	Pro	Tyr	Thr	Ala	Ser	Val	Val	Arg	Ser	Arg	Asn	Leu	Leu	Ile	
175				420					425					430			
177	TAT	GAG	GAA	CAC	GCT	CCC	CTG	GAA	GAG	TCG	TCG	AGC	CCA	GCC	TCA	ACC	1644
178	Tyr	Glu	Glu	His	Ala	Pro	Leu	Glu	Glu	Ser	Ser	Ser	Pro	Ala	Ser	Thr	
179				435				440					445				
181	GGG	GAG	CTG	GGG	GGC	TGG	GAC	AGA	GGC	TCC	TCT	CTG	CTC	AGG	GCA	TCG	1692
182	Gly	Glu	Leu	Gly	Gly	Trp	Asp	Arg	Gly	Ser	Ser	Leu	Leu	Arg	Ala	Ser	
183			450				455					460					
185	TCG	GGG	CTT	GAG	GCC	CTG	CCA	GAG	GTG	GAT	CTT	CCC	CAT	TTC	CTT	ATC	1740
186	Ser	Gly	Leu	Glu	Ala	Leu	Pro	Glu	Val	Asp	Leu	Pro	His	Phe	Leu	Ile	
187		465				470						475					
189	TCC	AAT	GAG	ACG	AGC	TTG	GAG	AAG	TCA	GTA	CTG	TTG	GAG	CTT	CAG	CAG	1788
190	Ser	Asn	Glu	Thr	Ser	Leu	Glu	Lys	Ser	Val	Leu	Leu	Glu	Leu	Gln	Gln	
191	480				485					490					495		
193	CAC	CTG	GTC	AGC	AGC	AAA	CTG	GAA	GGG	AAC	GAA	ACA	CTC	AAG	GTT	GTA	1836
194	His	Leu	Val	Ser	Ser	Lys	Leu	Glu	Gly	Asn	Glu	Thr	Leu	Lys	Val	Val	
195				500				505					510				
197	GAG	CTG	GAG	AGG	AGA	GTG	AAC	GCC	ACC	TTC	TGAGGAGAGC	TCCACCCTCC					1886

## RAW SEQUENCE LISTING

DATE: 02/14/2002

SEQUENCE APPLICATION: US/09/884,566A

TIME: 09:27:02

Input Set : A:\37497.txt

Output Set: N:\CRF3\02142002\I884566A.raw

```

198 Glu Leu Glu Arg Arg Val Asn Ala Thr Phe
199          515          520
201 TCTTCTCCTA ACTCCATCTG ATCGTCCTTT CAGTTTCCCC TTCACCACTG GATCTCGAGG 1946
203 AGGAGATGGC TAGTGTTACG GCTCGAGACA GGAGGCCAGC CCAGGGGGTT CTGCTTATGT 2006
205 GTCCCCGTGG CTCTCCACAA AAGGGAGCTA GCACCTCTCC ATCCCTTTCT CTTACTGCCA 2066
207 TTGGAAATTA TTTTAGGGCT GAGATAGGGG TGGAACGAGC AGGCTTGTTT CCACCAATAG 2126
209 TGCCAAGAAG AACTGCCTG ATTCTTCCCC GGGAGGAGTG ACTCCTCTGA AGAAGACATG 2186
211 ACTCATGTTT AGTTGAGACC CCAGACTCTA GCCACACACA TGCCACAGAC ATGCCAGGGA 2246
213 GTGGCAAAGC ACTGACTCCT GAGCTCCCTT CCTCACTAGG ACTCCAGTGT GACCCTGCAC 2306
215 TGAGAGGACC AAAGCGTCAT TGCAGTCTTC TCTCCACCCT GTACCCCGGA GTCCTGATTG 2366
217 GATGTCTGCA GAGGCAGATG GGGCTCCAC CATATTTTCA GGCCGCAAGT GCAATTCCTG 2426
219 AAGGCATCAG GCTCTTCTCT CCCAGGCTCT CCTGCCCACT GTGTTGTTTG TAGGACACCC 2486
221 CCACACCCAC TCATACACAG CCTGCATCTC CACAGGACAA TAGCTCTGTC TCCCTGGCCT 2546
223 CCCCTCCCA TTTGTAAATA GTATTTATTA GCTTGCTCAA GCTCCAGCT GGCCATAGTG 2606
225 AAAAGATTTT CCCTTTCAAC CAGCAAAGTC TTCTGTTGGC CTTTGGAACA GGAGAGTCCC 2666
227 CGGAATCTAG GACCCTAGTC TTTGTAAGTG ATGCCTTGTT TCCCCCTTT TCTTCTTTAA 2726
229 AATTGGGGAC CTATAACATC ATCGCTGTTG CGGAATCCAC TTAGGCATGT GTCCCCTGAT 2786
231 GGATGAATAC ATGGGAATGG TGGATACTGT CTTCTGACTC AGGCTCTAGG CTCCATGGCT 2846
233 TCCTCTCTCT GGTCCTGCCA CACAGAAGGA AAGCCCTGTC CAGGATAATG AGCGTTGCTG 2906
235 ACACCCTTGC TAGCTTGTC TGCCTACCTG CTTACCCAC TCCCTCACCT TCCTCCTTCC 2966
237 CTTCTGCCCT CCATCCACCT GCCTTAACTA ATTGGGGCTG GAGTTGGTCA TTTTTTGTAC 3026
239 ACCCACAGTG GTACCTTTTA CAGTCAGGTT TGGATACTTT GCAGCTCATC CAAAGAGACA 3086
241 TAACTAAACC CTAAACTCTT TTTTGTGTTG TGTTGTTGTT GTTTTTTTTT TTTATGATTA 3146
243 AAAAGTAAAA ATTGTAGTTT AAAAAAAAAA AAAAAAACT CGAG 3190

```

246 (2) INFORMATION FOR SEQ ID NO: 2:

248 (i) SEQUENCE CHARACTERISTICS:

249 (A) LENGTH: 521 amino acids

250 (B) TYPE: amino acid

251 (D) TOPOLOGY: linear

253 (ii) MOLECULE TYPE: protein

255 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

257 Met Glu Val Leu Glu Ser Gly Glu Gln Ser Val Leu Gln Trp Asp Arg
258 1          5          10          15
260 Lys Leu Ser Glu Leu Ser Glu Pro Gly Glu Thr Glu Ala Leu Met Tyr
261          20          25          30
263 His Thr His Phe Ser Glu Leu Leu Asp Glu Phe Ser Gln Asn Val Leu
264          35          40          45
266 Gly Gln Leu Leu Ser Asp Pro Phe Leu Ser Glu Lys Ser Glu Ser Met
267          50          55          60
269 Glu Val Glu Pro Ser Pro Thr Ser Pro Ala Pro Leu Ile Gln Ala Glu
270 65          70          75          80
272 His Ser Tyr Ser Leu Ser Glu Glu Pro Arg Thr Gln Ser Pro Phe Thr
273          85          90          95
275 His Ala Ala Thr Ser Asp Ser Phe Asn Asp Glu Glu Val Glu Ser Glu
276          100          105          110
278 Lys Trp Tyr Leu Ser Thr Glu Phe Pro Ser Ala Thr Ile Lys Lys Glu
279          115          120          125
281 Pro Ile Thr Glu Glu Gln Pro Pro Gly Leu Val Pro Ser Val Thr Leu
282          130          135          140

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/884,566A

DATE: 02/14/2002

TIME: 09:27:02

Input Set : A:\37497.txt

Output Set: N:\CRF3\02142002\I884566A.raw

```

284 Thr Ile Thr Ala Ile Ser Thr Pro Phe Glu Lys Glu Glu Ser Pro Leu
285 145 150 155 160
287 Asp Met Asn Ala Gly Gly Asp Ser Ser Cys Gln Thr Leu Ile Pro Lys
288 165 170 175
290 Ile Lys Leu Glu Pro His Glu Val Asp Gln Phe Leu Asn Phe Ser Pro
291 180 185 190
293 Lys Glu Ala Ser Val Asp Gln Leu His Leu Pro Pro Thr Pro Pro Ser
294 195 200 205
296 Ser His Ser Ser Asp Ser Glu Gly Ser Leu Ser Pro Asn Pro Arg Leu
297 210 215 220
299 His Pro Phe Ser Leu Ser Gln Ala His Ser Pro Val Arg Ala Met Pro
300 225 230 235 240
302 Arg Gly Pro Ser Ala Leu Ser Thr Ser Pro Leu Leu Thr Ala Pro His
303 245 250 255
305 Lys Leu Gln Gly Ser Gly Pro Leu Val Leu Thr Glu Glu Glu Lys Arg
306 260 265 270
308 Thr Leu Val Ala Glu Gly Tyr Pro Ile Pro Thr Lys Leu Pro Leu Thr
309 275 280 285
311 Lys Ser Glu Glu Lys Ala Leu Lys Lys Ile Arg Arg Lys Ile Lys Asn
312 290 295 300
314 Lys Ile Ser Ala Gln Glu Ser Arg Arg Lys Lys Lys Glu Tyr Met Asp
315 305 310 315 320
317 Ser Leu Glu Lys Lys Val Glu Ser Cys Ser Thr Glu Asn Leu Glu Leu
318 325 330 335
320 Arg Lys Lys Val Glu Val Leu Glu Asn Thr Asn Arg Thr Leu Leu Gln
321 340 345 350
323 Gln Leu Gln Lys Leu Gln Thr Leu Val Met Gly Lys Val Ser Arg Thr
324 355 360 365
326 Cys Lys Leu Ala Gly Thr Gln Thr Gly Thr Cys Leu Met Val Val Val
327 370 375 380
329 Leu Cys Phe Ala Val Ala Phe Gly Ser Phe Phe Gln Gly Tyr Gly Pro
330 385 390 395 400
332 Tyr Pro Ser Ala Thr Lys Met Ala Leu Pro Ser Gln His Pro Leu Ser
333 405 410 415
335 Glu Pro Tyr Thr Ala Ser Val Val Arg Ser Arg Asn Leu Leu Ile Tyr
336 420 425 430
338 Glu Glu His Ala Pro Leu Glu Glu Ser Ser Ser Pro Ala Ser Thr Gly
339 435 440 445
341 Glu Leu Gly Gly Trp Asp Arg Gly Ser Ser Leu Leu Arg Ala Ser Ser
342 450 455 460
344 Gly Leu Glu Ala Leu Pro Glu Val Asp Leu Pro His Phe Leu Ile Ser
345 465 470 475 480
347 Asn Glu Thr Ser Leu Glu Lys Ser Val Leu Leu Glu Leu Gln Gln His
348 485 490 495
350 Leu Val Ser Ser Lys Leu Glu Gly Asn Glu Thr Leu Lys Val Val Glu
351 500 505 510
353 Leu Glu Arg Arg Val Asn Ala Thr Phe
354 515 520
356 (2) INFORMATION FOR SEQ ID NO: 3:

```

VERIFICATION SUMMARY

DATE: 02/14/2002

PATENT APPLICATION: US/09/884,566A

TIME: 09:27:03

Input Set : A:\37497.txt

Output Set: N:\CRF3\02142002\I884566A.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:426 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]